

Základy využití korpusu v praxi **cjbb75**

Středa: 9.10-10.50 G13

5. 4. 2017

Jak lze v korpusech hledat doklady pro výzkum morfologie?

Vyhledávání podle morfologické značky – problém spolehlivosti desambiguace

Vyhledejte v korpusu SYN2010 všechny vokativy

Jak?

Hledat v korpusu

Korpus: syn2010 : --celý korpus--

Typ dotazu: CQL

CQL: vložit tag vložit "within" klávesnice předchozí dotazy
[tag="N...5.*"]

Specifikovat kontext ▼

Specifikovat dotaz podle metainformací ▼

Hledat

(můžete použít funkci **vložit tag**).

o své vůli . ŠÉFREDAKTOR Pletete páté přes deváté ,	mladíku /mladík/NNMS5-----A-----	. Subjektivně se jim mohlo zdát , že tam jít
Samozřejmě že jste měla na vybranou . " " Slečno	Coyneová /Coyneová/NNFS5-----A-----	, uzavřela bych obchod i se samotným ďáblem , aby
není patent na rozum . Mně spousta dětí říká paní	učitelko /učitelka/NNFS5-----A-----	nebo paní profesorko , já slyším na všechno . Proběh
" Hej , vy tam ! Okamžitě slezte dolů ,	chlape /chláp/NNMS5-----A-----	! " John Road se ještě nedotkl paluby , když
řekl : " Tohle by vás možná mohlo zajímat ,	doktore /doktor/NNMS5-----A-----	Wilde . " Ukázal na největší z displejů , připevněný
si perfektní oční linky . " Zapamatovala sis všechno ,	Judito /Judita/NNFS5-----A-----	? Proč nic neříkáš ? " zeptal se Teo po
nimi . . . Byl jsi u ní , strýčku	Dylane /Dylan/NNMS5-----A-----	. " " Ale copak jim to pomohlo ? " "
Donně Tumolové z Ridge , New York . Děkuji ,	dívenko /dívěnka/NNFS5-----A-----	! Osoby Morganovi z Ashby Hall Sir Robert - lord
, zavolala jsem : " Nedělej si starosti ,	mami /máma/NNFS5-----A---1-	. Jdu si lehnout a promluvíme si zítra . '
Ach , to jsem rád , že jsi přišel ,	bratře /bratr/NNMS5-----A-----	. Nutně , ale opravdu nutně potřebuji tvou pomoc !
ze vrat . Postrčí Vítko k maštalí Pospěš si ,	loudo /louda/NNMS5-----A-----	! 57 . obraz Krajina s jezerem a srubem Ráno
na lavičce . " To je v pořádku , pane	Jacobi /Jacobi/NNMS5-----A-----	, " řekl jsem , " díky , já už
, tak si najdu pořádnou práci . " " Pane	Simmons /Simmons/NNMS5-----A-----	, prosím vás , já ty peníze seženu . "
fidly . Jen aby se ti vůbec postavil , ty	srabe /srab/NNMS5-----A-----	. Abys dal vůbec dohromady jedno komický číslo , k
" Ať tě ani nenapadne lhát a něco zatloukat ,	slečinko /slečinka/NNFS5-----A-----	. To by ti u státního zástupce mohlo přijít hergot
zvučným hlasem . " Co pro vás můžu udělat ,	pane /pan/NNMS5-----A-----	Quinne ? " " Právě jsem hovořil s manželkou Patrick

Seznam lc (slovní tvary bez ohledu na velikost písma)

	<u>lc</u>	<u>Frekvence</u>	
1.	p/ n pane	9 305	■
2.	p/ n bože	3 902	■
3.	p/ n slečno	1 859	■
4.	p/ n doktore	1 340	■
5.	p/ n kapitáne	979	■
6.	p/ n otče	958	■
7.	p/ n chlapče	910	■
8.	p/ n kamaráde	903	■
9.	p/ n člověče	888	■
10.	p/ n maminko	695	■
11.	p/ n šéfe	650	■
12.	p/ n davide	618	■
13.	p/ n lásko	528	■
14.	p/ n tome	465	■
15.	p/ n bratře	465	■
16.	p/ n johne	436	■
17.	p/ n příteli	411	■
18.	p/ n holčičko	393	■
19.	p/ n kámo	391	■
20.	p/ n erno	379	■
21.	p/ n majore	346	■
22.	p/ n roberte	338	■
23.	p/ n sire	331	■
24.	p/ n paní	304	■
25.	p/ n mistře	304	■
26.	p/ n miláčku	302	■
27.	p/ n mylorde	295	■
28.	p/ n matko	293	■

Copak jsou v češtině vokativy jenom u substantiv, která nemají homonymní tvary vokativu?

zablokována cesta . Mnou . " Ty jim pomáháš ,	svině /svině/NNFS5-----A-----	? Tě zkrátím o dva metry , " vyslechnu si
tělem . " Kurva , nemůžu se hejbat ! Ty	svině /svině/NNFS5-----A-----	, tys mi přerazila páteř nebo co ! Utikej zpátky
to znovu požádal . " " Drž hubu , ty	svině /svině/NNFS5-----A-----	! " " A zas . " " Nech toho
to byl důvod Eviny existence . " Tak dělej ,	svině /svině/NNFS5-----A-----	, jdeme ! " Podal jí sklenici skotské , a
zezadu bezmocně řvali : " Ty svině ! Ty zatracená	svině /svině/NNFS5-----A-----	! " Mně to však utkvělo nejvíc , protože ti
a cpe do připravených antonů . Vy svině , vy	svině /svině/NNFP5-----A-----	, supí Sváta a všechno příkoří , které kdy prožil
Ó vy svině , vy si mě nepřejte . Vy	svině /svině/NNFP5-----A-----	. Vy si mě nepřejte . - Když otevřel vysoké
snad musela šlapat nebo co . A nemuselas ? Ty	svině /svině/NNFS5-----A-----	. Blázne . Ale ne . Jen se zkusím zachránit
důvod : Grulich mu prý předtím řekl " ty židovská	svině /svině/NNFS5-----A-----	" . V tomto kontextu vyznívá celý souboj podstatně jinak
Jága , pokřikovali na mě už z hledíště , ty	svině /svině/NNFS5-----A-----	! Komisař : Musel to být sugestivní výkon . Přítel
jsme příliš pozdě , pozdě o tři minuty . Ty	svině /svině/NNFS5-----A-----	, ty jsme dostali . Ale na ty chlapce ,
nesníme , ale ty s nimi taky šmelit nebudeš ,	svině /svině/NNFS5-----A-----	, " zaječela Kanello . " Šmelinář ! " "
pod krkem ! Řeknu : kde jsi byl , ty	svině /svině/NNFS5-----A-----	, utopenče , opustil's mě ? To je všechno ,

na srdci dnes ráno , Trevore ? " " Pane	učiteli /učitel/NNMS5-----A-----	St. Claire , jste ženatý ? " " Ne ,
. Mám o tebe strach . " " Můj nejmoudřejší	učiteli /učitel/NNMS5-----A-----	, zasel jsi do mého srdce tolik lásky , "
dovoleno s nimi vtipkovat . " Ať zazpívá , pane	učiteli /učitel/NNMS5-----A-----	! " říkají kluci a pan učitel Gouws švihne zápěstím
řekla : " Je nám to fakt líto , pane	učiteli /učitel/NNMS5-----A-----	, ale někdo zašantročil naši výstroj . " Pan Rittmann
jsi ho zase našel ? " " Ano , pane	učiteli /učitel/NNMS5-----A-----	. . . " Celý život mu byl vděčný za
uličkou za ním , a hned nato : " Pane	učiteli /učitel/NNMS5-----A-----	, teď nemám scénář . " Když se Megan blížila
, co máš ty ? " " Západ , pane	učiteli /učitel/NNMS5-----A-----	. " " Jak to , musíš tam mít sever
chlapecký hlas : Sim , to byla Svobodová , pane	učiteli /učitel/NNMS5-----A-----	, Svobodová . divčí hlas : Bonzáku , to ti
tak nesmyslného vzniknout . Vtom příroda vykřikla : - Pane	učiteli /učitel/NNMS5-----A-----	, pozor ! - Co se děje ? - Honem
Ať vezme svůj experiment s sebou . Ano , pane	učiteli /učitel/NNMS5-----A-----	. Na shledanou . - Na shledanou přírodu . Probudil
učej ! Jaký maj směšný školy a podívejte , pane	učiteli /učitel/NNMS5-----A-----	, na ty jejich učitele ! Jak se vytahou před
na mou rákosku ? - " No ne , pane	učiteli /učitel/NNMS5-----A-----	, " řekl Jacques , " tak vy jste si

Je desambiguace spolehlivá?

, že jsem tvůj muž . Můžu tě obejmout ,	anděli /anděl/NNMS5-----A-----	? Opera : Radši ne . Pusťte mě ! Pusť
a tisíc tisíců těch let i pravil Nejvyšší nejoddanějším svému	anděli /anděl/NNMS5-----A-----	, který Mu právě olizoval prsty u nohou : Táhni
. Starý pán : O to je to horší ,	anděli /anděl/NNMS5-----A-----	. Pohrdání je moje mateřský mlíko . Taky jsem k
samotné a sen kojení a čím mě lákáš , nebezpečný	anděli /anděl/NNMS5-----A-----	. Ty zrádně mlčíš a nedovolíš psát . Seru na
předchůdkyni Pohádky máje . Do věnování Józice napsal : "	anděli /anděl/NNMS5-----A-----	, obraze ideálu , zde první dech první mé žhavější
na rameni . . . Tak vypadá láska , můj	anděli /anděl/NNMS5-----A-----	! Na kterém oblaku přebýváš ? Italské nebe je čisté
mě vyvoláš . " A Évi oslovením " Ty můj	anděli /anděl/NNMS5-----A-----	" potvrzuje , že nemohou být znovu stvoření , že
Opera : Neslušnej ? Starý pán : Umiš vařit ,	anděli /anděl/NNMS5-----A-----	? Opera : Byla jsem dvakrát vdaná . Starý pán
se a byla klidná . Její radost byla ještě v	anděli /anděl/NNMS5-----A-----	strážném a její štěstí nebylo ještě v démonu . A
" andělec " , ale jak známó , i mezi	anděli /anděl/NNMS5-----A-----	jsou někteří padlí a jiní pracující v referátu smrti .
! Je to nechutný . . . Starý pán :	anděli /anděl/NNMS5-----A-----	. . . andíku ! Opera : (rozčileně)
a hodila ji do vzduchu . Padl orel . "	anděli /anděl/NNMS5-----A-----	strážný , prosím tě , stůj aspoň jednou při mně
sebou berete . Starý pán : Já děkuju tobě ,	anděli /anděl/NNMS5-----A-----	. Ani nevíš , jak jsem rád , že jsme

Ne tak docela.

Jak můžeme obejít chyby v desambiguaci?

Shrňme pozorování:

Seznam obsahoval převážně tvary vokativu, které v češtině nejsou homonymní s jinými pádovými formami. Které to jsou? Vokativ některých substantiv vzorů pán, hrad, muž a všech substantiv vzorů předseda a žena v sg. Všechna ostatní substantiva mají v češtině vokativ homonymní s nominativem (v plurálu). V sg. se homonymie liší u jednotlivých vzorů.

Pokusme se najít chyby v desambiguaci u tvarů vokativu substantiv maskulin životných.

Vzor pán: má vokativ na –e (není homonymní) a vokativ na –u, který se vyskytuje zpravidla u lemmat končících na veláru [kgh] s výjimkou substantiv *Bůh*, *člověk* a u substantiva *syn*.

Tvar na –u je ovšem homonymní, vyskytuje se totiž nejen ve vokativu, ale může stát také v lokále a dativu sg.

Vyhledejme v korpusu tvary vlku, zobrazme lemma a tag a podívejme se na chyby v desambiguaci

Hledat v korpusu

Korpus: : ▼

Typ dotazu: ▼

Slovní tvar:

▼

▼

to nedělal . " " Ty jsi ale hloupý , **vlku /v1k/NNMS3----**A---1- , vždyť my jsme něco jako bratr a sestra ,
, kreslil ptáky ohniváky a víly a princezny na šedém **vlku /v1k/NNMS6-----**A---1- a bohatýry zápasící s drakem ; vždycky to kreslil tím
lékem pořádně se proběhnout . " Taky by sis měl **vlku /v1k/NNMS3----**A---1- protáhnout nohy . Vsadím se s tebou , že bys
milionů) . Pákistánské akce proti al-Káidě připomínají přísloví o **vlku /v1k/NNMS6-----**A---1- , který se nažral , a koza zůstala celá .
musel pořád držet tuhle čepici višní ! " MY O **VLKU /v1k/NNMS6-----**A---1- ZA HUMNY Lámalý si ovce hlavu nad problémem , jak
máš . " (čínské přísloví : " my o **vlku /v1k/NNMS6-----**A---1- a vlk za dveřmi ") Nemůže jej opominout nikdo
a koza zůstala celá . Běžná je situace my o **vlku /v1k/NNMS6-----**A---1- a vlk za humny . " Vlk " jakožto opruzenina
1968 : . . . a přijely tanky aneb Beránka **vlku /v1k/NNMS3-----**A---1- se zachtělo evokuje text písničky Karla Kryla , bude přístupna
, podle toho , zda se podobali antickému mastífovi , **vlku /v1k/NNMS3-----**A---1- , chrtovi , jezevčíkovi nebo lišce . Pro zajímavost uvádíme
slyšeli hlas , který jim pravil : " Přijď , **vlku /v1k/NNMS5-----**A----- , půjč jim svou kůži , aby mohli rychle běžet
lesem a u cesty potká vlka : " Nazdar , **vlku /v1k/NNMS3-----**A---1- ! " povídá . " Ahoj , trdlo ! "
" " A nešlo by to přemýšlení trochu urychlit , **vlku /v1k/NNMS3-----**A---1- ? Však to možná ještě pamatuješ , mládí je netrpělivé
Medvěd mruče nahlédne do seznamu . " Ano , drahy **vlku /v1k/NNMS3-----**A---1- , a to dokonce už dnes večer . " Vlk
jenom já a ona . A když je řeč o **vlku /v1k/NNMS6-----**A---1- . . . " Na Sázkou je už jenom dnešek
vlk je ještě nezkušený . " Snad nejsi unavený , **vlku /v1k/NNMS3-----**A---1- ? " zeptala se liška . " To tedy jsem

Funkce rozhraní KonText:

Při práci s rozhraním kontext lze využít další funkci k označení řádků a jejich filtrování. Na obrázku je umístění. Volíme: Výběr řádků – základní/skupiny.

Výskytů: 205 | i.p.m. [?]: 0,08 vztaženo k celému syn | ARF [?]

Výběr řádků: skupiny (5 vybrané) 

#1 opus#1336

#1 opus#2003

#1 opus#26414

#1 opus#1336

#1 opus#1336

- opus#1336

- opus#40085

- opus#43365

Skupiny lze následně dále editovat/třídít/filtrovat/zrušit.

Adresu takto zpracovaného seznamu lze uložit a vždy se ke zpracovanému seznamu vrátit.

Úkol:

Vyhledejte v korpusu SYN2010 všechny tvary končící na *u* maskulin životných, jejichž lemma končí na *k*, *g*, *h*, *ch*.

Budeme používat jazyk **cql** (Corpus Query Language). Budeme používat zadání pomocí regulárních výrazů a logických operátorů pro konjunkci.

```
[tag="N.M.*" & lc=".*u" & lemma=".*[kgh]"]
```

Dotaz na každou pozici je v jazyce **cql** v hranatých `[]` závorkách.

Můžeme se ptát na lemma, tag, word, lc, pos, tedy na základní tvar, gramatické vlastnosti – kódované ve značce/tagu, slovní tvar, slovní tvar bez ohledu na to, jakými písmeny je napsán a na slovní druh (1. pozici ve značce – pos – **PART OF SPEECH**).

Do hranatých závorek napíšeme typ dotazu, tj. , na jaký typ se ptáme, bez mezery následuje znak ‚=‘ a v úvozovkách "" je slovní tvar/základní tvar/značka atd., na který/kterou se ptáme.

Pokud chceme kombinovat v rámci jednoho dotazu více podmínek, můžeme je spojit znaky pro logické operace (např. konjunkce, disjunkce, negace).

Uvnitř dotazu můžeme také pracovat s disjunkcemi.

Nyní slovně popíšu výše uvedený dotaz:

```
[tag="N.M.*" & lc=".*u" & lemma=".*[kgh]"]
```

vyhledej všechna substantiva maskulina : **tag="N.M.*"**

pro něž platí: **&**

že tvary bez ohledu na velikost písma končí na *u*: **lc=".*u"**

a pro něž platí: **&**

že základní tvar končí na *k*, nebo na *g*, nebo na *h* (tím zahrneme i případy, kdy končí na *ch*):
lemma=".*[kgh]"

([kgh] je v jazyce cql za disjunkce jednotlivých znaků/písmen)

Korpus: : --celý korpus-- ▾

Typ dotazu: ▾

CQL:

Specifikovat kontext ▾

Specifikovat dotaz podle metainformací ▾

Sledujeme chyby v desambiguaci

to nelíbí . A jí se to líbí . "	Oldřichu /Oldřich/NNMS3-----A---1-	, vážně byste ke mně nechtěl zajet na kávu ?
Major měl pleť proměněnou omlazovacími kúrami v cosi podobného narůžovělému	plastiku /plastik/NNMS3-----A---1-	. Byl poněkud otlý a o půl hlavy menší než
že by byly stačily pro tisíc bratří . A při	bohu /bůh/NNMS6-----A-----	živém , mezi tímto panovníkem , Chánem , a r
roce 1635 konali obyvatelé francouzské Remeše slavnostní procesí na poděkování	Bohu /bůh/NNMS3-----A-----	, že jejich město zbavil metly malomocenství
k ničemu , když k podobné události dojde doopravdy .	Člověku /člověk/NNMS3-----A-----	to přijde až líto , když pomyslí , kolik mozkové
brát jako obrazné . Mravní chování náleží k lásce k	Bohu /bůh/NNMS3-----A-----	a k bližnímu , pravda víry k poznání Boha a
. " Ahoj , Boffy ! " " Čau ,	drahoušku /drahoušek/NNMS3-----A---1-	! " Vděčně pospíchal dovnitř a zabouchl dveře
řekla tenkrát večer před svou konfirmací : Modlím se k	Bohu /bůh/NNMS3-----A-----	, abych byla užitečná . Vlastně jsem hrozně si
mu bylo zabaveno . STOPTIME MILANA TESAŘE POSPĚŠ SI ,	BOŘKU /Bořek/NNMS5-----A-----	" Dřevěná podlaha uprostřed přírody ? Co se t
Začal poletovat snít . " Znáte cíl cesty , pane	poručíku /poručík/NNMS5-----A-----	? " odhodlal se konečně strojevídce . " Neznár
jednoho stolu ? " Oba muži ještě pořád hovořili o	Bohu /bůh/NNMS6-----A-----	, když odešla do kuchyně připravit kávu . Ješt
obecná úprava je obsažena v občanském zákoníku a speciální v	zákoníku /zákoník/NNMS6-----A---1-	obchodním . Vztah ObčZ a ObchZ je v otázce
jsou stovky , tak jsme oba za vodou , ty	kluku /kluk/NNMS3-----A---1-	komediantská ! " " Češi , když se napijou ,
se po něm . " Kapitán uhodil vojáka . "	Hlupáku /hlupák/NNMS3-----A---1-	. " Ukázal na Borrice . " Tamten slyšel to
tříkrát nebo čtyřikrát . Prostě podle potřeby , " řekla	Blesku /blesk/NNMS3-----A---1-	. Strach z infekce Žádný zvláštní režim doktor
Když vám to udělá dobře , tak ano , pane	poručíku /poručík/NNMS5-----A-----	. " Moc mi toho namlouvat nemusíte , Clearb
a nebyl schopný ze sebe provodit kloudného věty . Pane	Irvingu /Irving/NNMS5-----A-----	, to , jak jste se tady objevil dnešní ráno
mého soukromého života . " " Tomu neunikneme nikdo ,	svobodníku /svobodník/NNMS5-----A-----	Benjamine , " řekl kapitán Frank . " Všechny d
mou naprostou insuficienci ! To je přece váš termin ,	soudruhu /soudruh/NNMS3-----A---1-	řediteli ! Ředitel však miní , že mi musí dát

Dotaz vylepšíme, abychom nevyhledali tvary *Bohu*, *bohu*, *člověku*, které nemohou být vokativy.

Jak?

Použijeme operátor negace a do dotazu zahrneme podmínku, že nechceme tvary lemmat *bůh* a *člověk*.

[tag="N.M.*" & lc=".*u" & lemma=".*[kgh]" & lemma!="(bůh)|(člověk)"]

přibližně stejně jako na Pýšu nebo na Kdepak , ty	ptáčku /ptáček/NNMS6-----A---1-	hnízdo máš ? Zimní přiběh o počátcích lyžování v
znepokojený zároveň . " Stalo se něco strašně smutného ,	miláčku /miláček/NNMS3-----A---1-	," řekla . " Kmotříček Pat dneska ráno zemřel
webové stránky ministerstva . Pravděpodobně chtěla psát o německém generálovi	Wolfgangu /Wolfgang/NNMS6-----A---1-	Schneiderhanovi , který je u nás na návštěvě , "
Tamní chlapi jsou samej vtípek . Pořád člověku říkaj "	náčelníku /náčelník/NNMS3-----A---1-	a mluvej o ohnivý vodě a tak . Mně je
od toho , co je za přírodou . Ve stejném	duchu /duch/NNMS6-----A---1-	byla slova paní Mostynové , když opět promluvila . "
" Ach , já si dávám velký pozor , soudruhu	plukovníku /plukovník/NNMS3-----A---1-	. Jsem lékař , vždycky pečlivě vážím slova . Když
milovala , milovala Ernsta Polaka . Ten je pro ni	pánbu /pánbůh/NNMS1-----A---2-	, pro něho shání kdejakou novinku , divadelní drb ,
jeden kočí s druhým , je - li kůň podoben	býku /býk/NNMS3-----A---1-	a vidí též červeně , a zatím mu kone uletli
utroubit . Sumec přistoupil k Holanďanovi . " Zalez do	stěhováku /stěhovák/NNMS3-----A---1-	, buřte , " oslovil ho temně , " a
cizí lidí do domu a tak řekneme Dobrou noc Sbohem	poutníku /poutník/NNMS3-----A---1-	a on jde někam o dům dál . . .
rámcí standardního postupu poskytla EK informaci o menší havárii v	Kršku /Kršek/NNMS6-----A---1-	všem členským státům . Bratislava prochází proměnou , chybí jen
a druh podnikání , vztahující se zpravidla k této osobě nebo	druhu /druh/NNMS3-----A---1-	podnikání . Dodatky nejsou povinnou součástí firmy fyzické osoby .
mnohém pravdu , " pokračovala Markétka v pro sebe neobvyklém	duchu /duch/NNMS6-----A---1-	. " Dokonce jsem ho poprosila , aby se ke
remízu v předkole LM Reprezentační obránce Radoslav Kováč získal fotbalistům	Spartaku /Spartak/NNMS3-----A---1-	Moskva nadějný výsledek pro odvetu 2 . předkola Ligy mistrů
blázen , namouduši , " svěřil se Wenck náčelníkovi štábu	plukovníku /plukovník/NNMS3-----A---1-	Reichhelmovi . " Třeba jim došly zásoby a potřebují se
a ve Zvolenu , která zásobovala celé střední Slovensko . Patřila	Čechu /Čech/NNMS3-----A---1-	Karlu Svobodovi . Jeho po vyhlášení Slovenského státu v březnu
Euforické 1 . . Svěží Orient , vůně s květem	hedvábníku /hedvábník/NNMS3-----A---1-	, Eau de Fleur de Soie , Kenzo , EdT
. Takže myslím , že z toho nic nebude ,	právníku /právník/NNMS3-----A---1-	z ženský prdele . Myslím , že budu spíš potřebovat
" Maminka mu říká Vendo , Vášo , Vašíku nebo	Venoušku /Venoušek/NNMS3-----A---1-	ale nikdy mu neříká Fanoušku . A on tvrdí

Všimněme si, kde se nacházejí chyby.

Pokuste se formulovat nějakou vlastnost, jíž se liší (morfosyntakticky) tvar vokativu od tvaru dativu a lokálu.

Např.: před tvarem lokálu/dativu je jméno ve stejném pádě, nebo předložka pojící se se stejným pádem.

Nebo: Tvary vokativu bývají v češtině odděleny čárkou nebo koncovou interpunkcí.

Nebo: před tvarem vokativu se bezprostředně nachází jiný tvar ve vokativu (*pane poručíku*, *můj miláčku*).

Nebo: Před tvarem vokativu stojí osobní zájmeno *ty* (*ty hlupáku*).

Na základě takovýchto vlastností můžeme „odfiltrovat“ řádky, které s velkou mírou pravděpodobnosti budou obsahovat tvary se správnou značkou.

Z následujícího seznamu lemmat vybereme řádky, na nichž je tvar *miláčku*.

Celkem: 1614 (33 str.)

	lemma	lc	Frekvence	
1.	p/ n miláček	miláčku	1 557	
2.	p/ n duch	duchu	1 507	
3.	p/ n poručík	poručíku	645	
4.	p/ n drahoušek	drahoušku	633	
5.	p/ n soudruh	soudruhu	445	
6.	p/ n hoch	hochu	376	
7.	p/ n tatínek	tatínku	331	
8.	p/ n strýček	strýčku	307	
9.	p/ n plukovník	plukovníku	282	
10.	p/ n František	františku	264	
11.	p/ n Jack	jacku	260	
12.	p/ n synek	synku	224	
13.	p/ n Hrádek	hrádku	181	
14.	p/ n Blesk	blesku	171	
15.	p/ n Zdeněk	zdeňku	165	
16.	p/ n Jindřich	jindřichu	162	
17.	p/ n Patrick	patricku	145	
18.	p/ n klokánek	klokánku	144	
19.	p/ n pánbůh	pámbu	134	
20.	p/ n Bismarck	bismarcku	127	
21.	p/ n dědeček	dědečku	123	
22.	p/ n Mark	marku	121	
23.	p/ n Jeník	jeníku	120	

znepokojený zároveň . " Stalo se něco strašně smutného ,	miláčku /miláček/NNMS3-----A---1-	, " řekla . " Kmotříček Pat dneska ráno zemřel
výrazně snižují riziko infarktu . Takže vaše heslo zní -	miláčku /miláček/NNMS3-----A---1-	, nenech se prosit a pojd' mi zachránit život !
dal . Ale je to všechno , co mám ,	miláčku /miláček/NNMS3-----A---1-	. Převážím to z jedné stanice na druhou . Každá
k Jeffovi . " Hodně štěstí v novém roce ,	miláčku /miláček/NNMS6-----A---1-	. " " Tobě taky , Sáro . " Usnuli
jsem přece neřek Jo mám velké zadek a tlustý stehna	Miláčku /miláček/NNMS3-----A---1-	proboha vůbec jsem o tvým zadku nemluvil Říkals že mi
je ? " Další květináč s pelargonii mě ,	miláčku /miláček/NNMS3-----A---1-	, praštil do kříže , skutálel se mi na prdel
do jeho zrudlé tváře . " To nebylo špatný ,	miláčku /miláček/NNMS3-----A---1-	, " zašeptala , když mu tiskla vášnivý polibek na
? Co bys asi tak řekl ? Skvělý nápad ,	miláčku /miláček/NNMS3-----A---1-	, ale vyber si radši jinou učitelku ? Marna je
mi nehty až do masa . " Už jsem ,	miláčku /miláček/NNMS3-----A---1-	, už jsem ! " " Já taky , "
zeď a nesesul se k zemi . - Ó ,	miláčku /miláček/NNMS3-----A---1-	, vzlykala Terry a věšela se Bernhardovi kolem krku .
do pusy . " Do háje , jsou skvělý ,	miláčku /miláček/NNMS3-----A---1-	! OCHUTNEJ ! " Joyce nabodla jednu terčovku na vidličku
sebe , protože vy mě přece milujete . " "	Miláčku /miláček/NNMS3-----A---1-	- " Catherine natáhla ruku , aby ho umčela .
, šest kuřat , vypiju láhev šampaňského a šup ,	miláčku /miláček/NNMS3-----A---1-	, jdeme na to . " " A když se
kulatým obličejem a velkýma očima . Hoskíns řekl : "	Miláčku /miláček/NNMS3-----A---1-	, to je slečna Fellowesová , ošetřovatelka , která se
jsem a udělala na něj oči . " Dobrá ,	miláčku /miláček/NNMS3-----A---1-	, budu ti něco vyprávět ! " libnul mě lehce
grand-père ? " škemrala unaveně Sophie . " Brzy ,	miláčku /miláček/NNMS3-----A---1-	, velmi brzy . " Jeho hlas byl plný melancholie
! " " Udržet tlak ? " " Víš ,	miláčku /miláček/NNMS5-----A---1-	, musíme tyrany držet pod krkem výhrůzkami . Samozřejmě to
Harry jen zíral . " Zapomněla jsem ti říct ,	miláčku /miláček/NNMS3-----A---1-	, " vyblekotala jsem , " tohle byl ten případ
co to je za úkol ? " " Kuny ,	miláčku /miláček/NNMS3-----A---1-	, děkuju ti , " vyhrkla a zcela nečekaně se
příliš mladá . Ale my víme své , vid' ,	miláčku /miláček/NNMS3-----A---1-	? " Důvěřivě přikývla . " Elizabeth , mám od

Všimněme si, jak vypadá okolí chybně označkových tvarů.

Tvary vokativu bývají v češtině odděleny čárkou, nebo koncovou interpunkcí.

Jak odstraníme řádky, na nichž budou v pozici KWIC tvary oddělené interpunkcí (<, miláčku, > nebo <, miláčku.> nebo <. Miláčku,>)?

Použijeme pozitivní filtr a zadáme, že chceme vyhledat v intervalu -1,1 příslušné interpunkční znaky:

[lemma="[\\.:!?!"]]

Filtr konkordancí

Filtr: pozitivní negativní

Vybraný token: první poslední

Rozsah hledání: od -1 do 1 včetně KWIC

Typ dotazu: CQL

CQL:
[lemma="[\\.:!?!"]]

Získáme seznam 1507 řádků, které patrně budou obsahovat chyby. Pro jistotu je projdeme. Můžeme se také podívat, zda některé tvary nejsou označovány správně.

Celkem: 3 (1 str.)

	lc	tag	Frekvence	
1.	p/n	miláčku	NNMS3-----A---1-	1 143
2.	p/n	miláčku	NNMS5-----A-----	300
3.	p/n	miláčku	NNMS6-----A---1-	64

Všimněme si, že podíl lokálů je menší, než podíl dativů. A zkontrolujme, zda existují řádky, na nichž by se v levém kontextu vyskytoval tvar v dativu.

Filtr: pozitivní negativní

Vybraný token: první poslední

Rozsah hledání: od do včetně KWIC

Typ dotazu:

CQL:

. " Odmícela se a pak : " <i>Petře</i> , miláčku /miláček/NNMS3-----A---1- . . . " Co je ? " "
Ach , stisknout článek tvého malíčku levého . <i>Miláčku</i> , miláčku /miláček/NNMS3-----A---1- , něco z toho přeci zůstane . Vždycky mít nabroušenou
bušivým , napjatým zvukem varhan . Chtěl bych <i>ti říkat</i> miláčku /miláček/NNMS3-----A---1- . Snažila se vykřiknout , ale hlas jí uvízl v
a Randy si začal pobrukovat " Chtěl bych <i>ti říkat</i> miláčku /miláček/NNMS3-----A---1- . " Claire při tom cítila žaludek až v krku
bude po večerech zpívat . - Ale , <i>Izáku</i> , miláčku /miláček/NNMS3-----A---1- , ať je to dítě jenom naše a ať na
Zeptáme se , jestli DJ nemá Chtěl bych <i>ti říkat</i> miláčku /miláček/NNMS3-----A---1- . " řekl . Nevěděla , jestli jí chce dráždit
Nuly gaussů vyslal směřované volání . " Skvrnitý <i>vršku</i> , miláčku /miláček/NNMS3-----A---1- ! Přijď sem , prosím tě . Máme hosty .
, jen na papíru . Podal ho Danymu <i>Polnauerovi</i> . Miláčku /miláček/NNMS3-----A---1- . Nemůžu to ani vypovědět , jak jsem šťastná ,
Nepodívala se na něj . " Co je <i>ti</i> , miláčku /miláček/NNMS3-----A---1- ? Ublížil jsem ti ? " " Ne , "
" Umírám před ní hrůzou . " <i>Beth</i> , miláčku /miláček/NNMS3-----A---1- , ale tvoje matka se nic nedozví ! " Konejšivě
, a já mu to dávala , stejně jako <i>tobě</i> miláčku /miláček/NNMS3-----A---1- . Dnešní svět je šílený . Proč jsi nemohl stát
, a . . . Ano ! Chtěl bych <i>ti říkat</i> miláčku /miláček/NNMS3-----A---1- . Slyšela jsem to z hudební skříně a málem jsem
ani správně hrát , ty - " <i>Luku</i> , miláčku /miláček/NNMS3-----A---1- - " Mozek žádný , kozy žádný , zadek
" Já myslím , že Lily ještě obstarává <i>šití</i> , miláčku /miláček/NNMS3-----A---1- . " " Kdepak , madam , to Růžena ,
polbí na ústa . " Všechno nejlepší k <i>výročí</i> , miláčku /miláček/NNMS3-----A---1- . " " Ale výročí máme až - - "
, všeobjímající lásku ? " " Jenom k <i>tobě</i> , miláčku /miláček/NNMS3-----A---1- , " odpověděla . " Tobě se nikdy nestává ,
. Ale fňukat nehodlal . " No tak <i>ti</i> , miláčku /miláček/NNMS3-----A---1- , děkuju , " řekl , automaticky šek složil a
snad úplně vymažeš mozek ! " " Děkuju <i>ti</i> , miláčku /miláček/NNMS3-----A---1- . Nebyl to můj nejlepší pokus , ale byl jsem
opřela si hlavu o jeho tvář . " <i>Miláčku</i> , miláčku /miláček/NNMS3-----A---1- , " hladil ji po vlasech . " Všechno bude
mluvit chlácholivé a láskyplně . Říkal : " <i>Rose</i> , miláčku /miláček/NNMS3-----A---1- , podívej se , ty taháš za ocas moc divokou
, omlouvala ho , a včera v noci <i>mi</i> řekl miláčku /miláček/NNMS3-----A---1- . Ohlédla se na zástup koní a hřibát , jež
nezná , máš to mnohem těžší . Řeknu <i>ti</i> , miláčku /miláček/NNMS3-----A---1- , aby ses v naší branži prosadila , potřebuješ všechno
Zapíchlá mu prst do hrudi . " A neříkejte <i>mi</i> miláčku /miláček/NNMS3-----A---1- . Jmenuju se Peabodyová . Strážník Peabodyová , a nepotřebuju
jiným Nožikům . A Pochva povídá Nožikovi : <i>Nožiku</i> , miláčku /miláček/NNMS3-----A---1- , vy jste pěkný kurevník ! Každý den měníte pochvu

Všimněme si, že tento aspekt je pro desambiguaci nerelevantní. Jména v dativu nerozvíjejí tvar substantiva v pozici KWIC (*svému miláčku*), ani s ním nejsou koordinovány (*Rostovi, miláčku svému, Vlastíkovi, drahoušku mému, ...*).

Nicméně můžeme se pokusit ještě více specifikovat dotaz.

Vrátíme se zpět a zadání zpřesníme:

Filtr konkordancí

Filtr: pozitivní negativní

Vybraný token: první poslední

Rozsah hledání: od -2 do -1 včetně KWIC

Typ dotazu: CQL

CQL:

[vložit tag](#) [vložit "within"](#) [klávesnice](#) [předchozí dotazy](#)

```
[tag="( [NAP] ..S3. * ) | (R...3. * )" & tag!="P[7HPY].S3. *"]
```

Filtrovat konkordanci

. " Odmělčela se a pak : " <i>Petře</i> ,	miláčku /miláček/NNMS3-----A---1-	... " Co je ? " "
Ach , stisknout článek tvého malíčku levého . <i>Miláčku</i> ,	miláčku /miláček/NNMS3-----A---1-	, něco z toho přeci zůstane . Vždycky mít nabroušenou
bude po večerech zpívat . - Ale , <i>Izáku</i> ,	miláčku /miláček/NNMS3-----A---1-	, ať je to dítě jenom naše a ať na
Nuly gaussů vyslal směřované volání . " Skvrnitý <i>vršku</i> ,	miláčku /miláček/NNMS3-----A---1-	! Přijď sem , prosím tě . Máme hosty .
, jen na papíru . Podal ho Danymu <i>Polnauerovi</i> .	Miláčku /miláček/NNMS3-----A---1-	. Nemůžu to ani vypovědět , jak jsem šťastná ,
" Umírám před ní hrůzou . " " <i>Beth</i> ,	miláčku /miláček/NNMS3-----A---1-	, ale tvoje matka se nic nedozví ! " Konejšivě
ani správně hrát , ty - " " <i>Luku</i> ,	miláčku /miláček/NNMS3-----A---1-	- " Mozek žádnéj , kozy žádný , zadek
" Já myslím , že Lily ještě obstarává <i>šití</i> ,	miláčku /miláček/NNMS3-----A---1-	. " " Kdepak , madam , to Růžena ,
polbil na ústa . " Všechno nejlepší k <i>výročí</i> ,	miláčku /miláček/NNMS3-----A---1-	. " " Ale výročí máme až - - "
opřela si hlavu o jeho tvář . " <i>Miláčku</i> ,	miláčku /miláček/NNMS3-----A---1-	, " hladil ji po vlasech . " Všechno bude
mluvit chlácholivě a láskyplně . Říkal : " <i>Rose</i> ,	miláčku /miláček/NNMS3-----A---1-	, podívej se , ty taháš za ocas moc divokou
jiným Nožíkům . A Pochva povídá Nožíkovi : <i>Nožíku</i> ,	miláčku /miláček/NNMS3-----A---1-	, vy jste pěkný kurevník ! Každý den měníte pochvu
Arlene ! " Křečovitě se roztřásla . " <i>Miláčku</i> ,	miláčku /miláček/NNMS3-----A---1-	, miláčku - vyvádíš naprosto zbytečně . Moje pokusy nezahnují
němu a chytila ho kolem krku . " <i>Miláčku</i> ,	miláčku /miláček/NNMS3-----A---1-	, ono se to zlepší , uvidíš . " Týden
bylo ticho a pak promluvil Parrish : " <i>Sally</i> ,	miláčku /miláček/NNMS3-----A---1-	, tímhle přece nic - " Otočila se k němu
já " se bouřilo pod dekou . " <i>Miláčku</i> ,	miláčku /miláček/NNMS3-----A---1-	! " Ruka sjela z prsů na tvrdé rovné břicho
Janet a já jí ho představil . " <i>Janet</i> ,	miláčku /miláček/NNMS3-----A---1-	, tohle je Bernard Westcott . Za starých časů to
Haló ! Výborně . Nefunguje to . (<i>Gréte</i>)	Miláčku /miláček/NNMS3-----A---1-	, rychle mizíme . Prcek : Jindro , prosím tě
žádné holce jsem necítil něco takového jako k <i>Tobě</i> .	Miláčku /miláček/NNMS3-----A---1-	, měj se moc a moc krásně a taky mi
Pak Anne-Marie rychle a naléhavě řekla : " <i>Jayí</i> ,	miláčku /miláček/NNMS3-----A---1-	, poslouchej . Použij sedmé nastavení . Sedmé . Slyšíš
Ano , - ano , ach ano - <i>miláčku</i> ,	miláčku /miláček/NNMS3-----A---1-	! Já se o tebe budu tak krásně starat ,
nezachrání . Všechno je marný ! " " <i>Miláčku</i> ,	miláčku /miláček/NNMS3-----A---1-	, tohle nesmíš říkat , " zašepřel Belford a položil
a pojednou byla schopna promluvit : " " <i>Hansi</i> ,	miláčku /miláček/NNMS3-----A---1-	, vypadáš tak unaveně . ' ' ' ' Všichni
které se jí zadržlo v hrdle . " <i>Králíčku</i> ,	miláčku /miláček/NNMS3-----A---1-	můj , " vzdychla . Pak už jí přemohly slzy

Všimněme si, jak chyby v disambiguaci vzájemně souvisejí:

Frekvenční distribuce

Frekvenční limit: 0

Nastavit limit

Celkem: 21 (1 str.)

	lemma	lc	tag	Frekvence
1.	p/ n miláček	miláčku	NNMS3-----A---1-	9
2.	p/ n šití	šití	NNNS3-----A-----	1
3.	p/ n výročí	výročí	NNNS3-----A-----	1
4.	p/ n vršek	vršku	NNIS3-----A-----	1
5.	p/ n rosa	rose	NNFS3-----A-----	1
6.	p/ n petra	petře	NNFS3-----A-----	1
7.	p/ n nožík	nožíku	NNIS3-----A-----	1
8.	p/ n luk	luku	NNIS3-----A-----	1
9.	p/ n králíček	králíčku	NNMS3-----A---1-	1
10.	p/ n Toba	tobě	NNFS3-----A-----	1
11.	p/ n Sally	sally	NNFS3-----A-----	1
12.	p/ n Polnauer	polnauerovi	NNMS3-----A-----	1
13.	p/ n Paula	paule	NNFS3-----A-----	1
14.	p/ n Noah	noahu	NNMS3-----A---1-	1

Vrátíme se a odfiltrujeme všech 1507 řádků z 1557 a 50 řádků uložíme, projdeme a označíme chyby v lemmatizaci.

Správně tagované jsou pouze 3 doklady, ve všech ostatních případech jde o chybně otagované vokativy.

<input type="checkbox"/>	doc#1409,Ženy, Havel, hygiena	jsem přece nelek Jo mám velké zadek a tlustý stehna	Miláčku /miláček/NNMS3-----A---1-	proboha vůbec jsem o tvým zadku nemluvil říkals že mi
<input type="checkbox"/>	doc#670,Nenasytá	sebe, protože vy mě přece milujete ."	Miláčku /miláček/NNMS3-----A---1-	" Catherine natáhla ruku , aby ho umlčela .
<input type="checkbox"/>	doc#139,Chladnokrevně	hlavní roli kocour - šlo o ten záhadný skon Nancyina	miláčku /miláček/NNMS3-----A---1-	Boobse , kterého podle záznamu pořízeného čtrnáct dní před vlastní
<input type="checkbox"/>	doc#15516,Štěstí začíná uvnitř	byť štěstím bez sebe , když mu někdo řekne "	miláčku /miláček/NNMS3-----A---1-	" , zatímco jiný se už jen přití předstává
<input checked="" type="checkbox"/>	doc#78726,Lidové noviny, 5. 5. 2007	nevi , jestli dojde . " Marhoulova rodina říká svému	miláčku /miláček/NNMS3-----A---1-	na čtyřech kolech Winston . Samozřejmě po Winstonu Churchillovi .
<input type="checkbox"/>	doc#807,Panna v Paříži	přítiská si konečky jejich chladných prstů ke rtům . "	Miláčku /miláček/NNMS3-----A---1-	" vydechl . " Přijel jsem za tebou , abych
<input type="checkbox"/>	doc#93,Železná brána	křídly ve voliéře - věčná prázdnota - ach noc	miláčku /miláček/NNMS3-----A---1-	vyved mě z toho žaláře skrze dveře vedoucí od smutku
<input type="checkbox"/>	doc#178,Peško v ráji	ze zad . " " A vombat ? " "	Miláčku /miláček/NNMS3-----A---1-	bohů , splnil jsi úkol tak dokonale , že ani
<input type="checkbox"/>	doc#36,Spiknutí škorpiónů	tim parchantem kvít . Vytáhl mě odtamtud ! " "	Miláčku /miláček/NNMS3-----A---1-	můj ! " zvolala majorka nevojensky , pustila kapitána a
<input type="checkbox"/>	doc#36,Spiknutí škorpiónů	se . V hlavě mu pořád vrtalo prostě oslovení "	miláčku /miláček/NNMS3-----A---1-	" , kvůli němuž se choval jako hlupák . jako
<input type="checkbox"/>	doc#5828,Na shledanou zítra (část 1.)	? Těhotná Marie Mačká se někde v metru Bojím se	miláčku /miláček/NNMS3-----A---1-	Mám strach že potratí Hudbu napsal Petr Skoumal a přejmenoval
<input type="checkbox"/>	doc#814,Osudný zvrst	, budou zdvořilí , budou se usmívat a říkat si	miláčku /miláček/NNMS3-----A---1-	častěji , než může být upřímné . Ale v ten
<input type="checkbox"/>	doc#912,Posedlá romantikou	celý večer a zároveň poprvé , co mi řekl "	miláčku /miláček/NNMS3-----A---1-	" . Co se děje ? Že by zářil na
<input type="checkbox"/>	doc#759,Linkie krásy	radostí . V duchu se upnul k tomu nebekanému "	miláčku /miláček/NNMS3-----A---1-	" , které ho hrálo a zamotalo mu hlavu stejně
<input type="checkbox"/>	doc#528,Starí muži o půlnoci	" " Pokud se to ještě zhorší - " "	Miláčku /miláček/NNMS3-----A---1-	" - " měli bychom si sem vzít noční
<input type="checkbox"/>	doc#354,Zapomenuté vzpomínky	se z ní linou melodie " Chtěl bych ti říkat	miláčku /miláček/NNMS3-----A---1-	" . Byla to jen nevinná písnička , ale hned
<input type="checkbox"/>	doc#1457,Vítejte v opičárně	na něj , obličej bílý jako obláček mlhy . "	Miláčku /miláček/NNMS3-----A---1-	- jsi v pořádku ? " zeptala se tise .
<input type="checkbox"/>	doc#36,Spiknutí škorpiónů	letectva , jež se baví s podřízeným . Slovíko "	miláčku /miláček/NNMS3-----A---1-	" se mělo říkat tise - bud pročitěné , anebo
<input type="checkbox"/>	doc#533,Světlo dne	vybral . Vyletelo mi z úst jako ptáček : "	miláčku /miláček/NNMS3-----A---1-	" . Takhle jsem Helenu ještě nikdy neoslovil . V
<input type="checkbox"/>	doc#5819,Zuřený	nebo mírová hlohovice Tak letím přes pole a lesy můj	miláčku /miláček/NNMS3-----A---1-	kde jsi ? Proletáři jsou barbaři zato kněžna ta je
<input type="checkbox"/>	doc#1409,Ženy, Havel, hygiena	atomovka a ty jsi mi natočil hlavu přímo k výbuchu	Miláčku /miláček/NNMS3-----A---1-	byl to jen nákladák Nemám nátek obličej ? Lásko to
<input type="checkbox"/>	doc#8,Ruská sekce	predal Igorovi . I to je docela normální . "	Miláčku /miláček/NNMS3-----A---1-	Káto . " - takhle tedy normálně nezaciná , to
<input checked="" type="checkbox"/>	doc#626,Strom smíření	Jemma vdaná (vždycky o něm mluvíval jako o "	miláčku /miláček/NNMS3-----A---1-	Cyrusovi ") , než padl v korejské válce .

https://kontext.korpus.cz/view?q=~D4I02CI0;fromp=1;corpname=omezeni%2Fsyn2010&viewmode=kwic&pagesize=40&attrs=word%2Clemma%2Ctag&attr_allpos=kw&ctx_attrs=word&structs=p%2Cg%2Cerr%2Ccorr&refs=doc%2C%3Dopus.nazev

ÚKOL na 12. 4. 2017

Ze seznamu

Celkem: 1614 (33 str.)

	<u>lemma</u>	<u>lc</u>	<u>Frekvence</u>	
1.	p/n miláček	miláčku	1 557	
2.	p/n duch	duchu	1 507	
3.	p/n poručík	poručíku	645	
4.	p/n drahoušek	drahoušku	633	
5.	p/n soudruh	soudruhu	445	
6.	p/n hoch	hochu	376	
7.	p/n tatínek	tatínku	331	
8.	p/n strýček	strýčku	307	
9.	p/n plukovník	plukovníku	282	
10.	p/n František	františku	264	
11.	p/n Jack	jacku	260	
12.	p/n synek	synku	224	
13.	p/n Hrádek	hrádku	181	
14.	p/n Blesk	blesku	171	
15.	p/n Zdeněk	zdeňku	165	
16.	p/n Jindřich	jindřichu	162	
17.	p/n Patrick	patricku	145	
18.	p/n klokánek	klokánku	144	
19.	p/n pámbůh	pámbu	134	
20.	p/n Bismarck	bismarcku	127	
21.	p/n dědeček	dědečku	123	
22.	p/n Mark	marku	121	
23.	p/n Jeník	jeníku	120	

si každý student vybere dva tvary (začneme s tvarem *duchu* a budeme pokračovat).
Odstraní případy mezi interpunkcí a zbytek konkordančních řádků projde a vyznačí chyby v lemmatizaci (uloží seznam konkordančních řádků a vyznačí je). Seznam s vyznačenými chybami odevzdá do 18. 4. 2017.